

8.1.1 OPTIMUM CODING TECHNIQUES FOR MST RADARS

Michael P. Sulzer

Arecibo Observatory
Box 995
Arecibo, Puerto Rico

and

Ronald F. Woodman

Instituto Geofisico del Peru
Apartado 3747
Lima, Peru

INTRODUCTION

The optimum coding technique for MST radars is that which gives the lowest possible sidelobes in practice and can be implemented without too much computing power. Coding techniques are described in FARLEY (1985). The best technique, in theory, is the complementary code pair. Coherent integration can be used to reduce the size of the data set, and so the amount of computation is not excessive. The sidelobes are zero in theory, but when errors induced by imperfections in the modulation of the transmitter are significant, the quasi-complementary set gives better results (SULZER and WOODMAN, 1984). However, this technique requires an extraordinary amount of computation. We discuss here a technique mentioned briefly in FARLEY (1985), but not fully developed and in general use. This is decoding by means of a filter which is not matched to the transmitted waveform, in order to reduce sidelobes below the level obtained with a matched filter. This is the first part of the technique discussed here; the second part consists of measuring the transmitted waveform and using it as the basis for the decoding filter, thus reducing errors due to imperfections in the transmitter. There are two limitations to this technique. The first is a small loss in signal-to-noise ratio, which usually is not significant. The second problem is related to incomplete information received at the lowest ranges. Appendix A shows a technique for handling this problem. Finally, we show that the use of complementary codes on transmission and non-matched decoding gives the lowest possible sidelobe level and the minimum loss in SNR due to the mismatch.

THE CODING-DECODING PROCESS

A model of the coding-decoding process starts with a square pulse of length t where t corresponds to the desired range resolution and the square pulse is described by $h_{sq}(t)$, since the pulse may be thought of as the response of a filter to an impulse, and thus is identified by the impulse response of this filter. This square pulse is what we would like to transmit if we had sufficient peak power. If no coding is done, the received signal is passed through a matched filter and is given by

$$s(t) = h_{sq}(t) * h_{sq}(-t) \quad (1)$$

The impulse response of the matched filter is just the flip of that of the transmitted signal, or in the frequency domain, the amplitude responses are the same and the phases are additive inverses.

If we use a phase code, then

$$s(t) = h_{sq}(t) * h_c(t) * h_{sq}(-t) * h_{dc}(t) \quad (2)$$

$h_c(t)$ is the impulse response of the coding filter. For a binary phase code $h_c(t)$ is a sequence of positive and negative impulses. For normal decoding $h_{dc}(t)$ is the flip of $h_c(t)$. For a perfect code $h_c(t)*h_c(-t)$ is an impulse. Imperfect codes will give sidelobes.

Consider the function $h_{dc}(t)$ for an arbitrary code such that

$$h_c(t)*h_{dc}(t) = [\text{impulse}] \quad (3)$$

This is the decoding function which eliminates all sidelobes. It exists for most codes, and it is calculated from the Fourier transform of $h_c(t)$:

$$H_c(f) = A_c(f)e^{j[\phi(f)]} \quad (4)$$

Then the impulse response of the decoding filter with no sidelobes is

$$h_{dc}(t) = F^{-1}[(1/A_c(f))e^{j[-\phi(f)]}] \quad (5)$$

As long as $A_c(f)$ has no zeros the inverse exists. The perfect code has $A_c(f)$ equal to a constant, and requires no amplitude correction at all. For good codes, the amplitude function is nearly constant and thus, the inverse exists and varies little as a function of frequency. The effect of the inverse amplitude filter is to pass more random noise than in normal decoding. This is one cost of eliminating the sidelobes, one that is a function of how good the code is. With a good code such as the 13-bit Barker code, the loss in signal to noise ratio is about .25 dB, hardly significant. A randomly selected code might lose several dB. A second problem is that $h_{dc}(t)$ is infinitely long and thus can never be used exactly for deconvolution. Sometimes this does not matter, and there are techniques for minimizing the effect when it is important that sidelobes be kept very small at very close ranges.

EXPERIMENTAL RESULTS

The results of various types of decoding are shown in Figures 1 through 4. These consist of the transmitted 430-MHz signal and the received signal covering a total time period of 256 microsec. The transmitted signal was the output of a probe in the waveguide; the received signal consisted of ground clutter and atmospheric scatter. The figures show power versus range, and it is the ground clutter which is the dominant signal. The two signals were added at the i.f. (30 MHz) level and thus passed through a common 500-kHz Gaussian filter. The transmitter was coded with a 13-baud Barker code with 2 microsec baud length. The sampling rate was also 2 microsec. The response of the sampled transmitted waveform to the decoding process is called the system function, since it shows the response of the receiver and decoder to a very narrow target.

Figure 1 shows the power versus range obtained when the transmitted and received signals are decoded with the Barker code. The main lobe of the decoded transmitter signal is broadened by the 500-kHz. Gaussian filter and sidelobes are visible both before and after the main lobe. The sidelobes before the main lobe are very close to the expected -22 dB level. The sidelobes following the main lobe are quite different. We shall not discuss the generating mechanism of these sidelobes except to say that both finite bandwidth and nonlinearities are involved, since physical filters can only affect the signal at later times.

Figure 2 shows the same data decoded with the inverse of the Barker code. The sidelobes to the left of the main lobe have been considerably reduced, since the sidelobes due to the code have been removed, while those that are

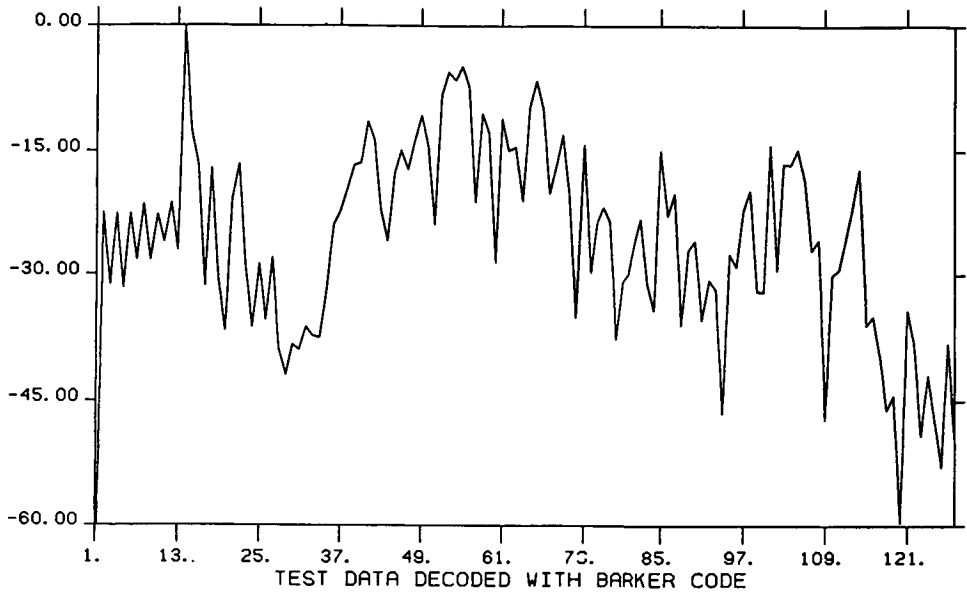


Figure 1.

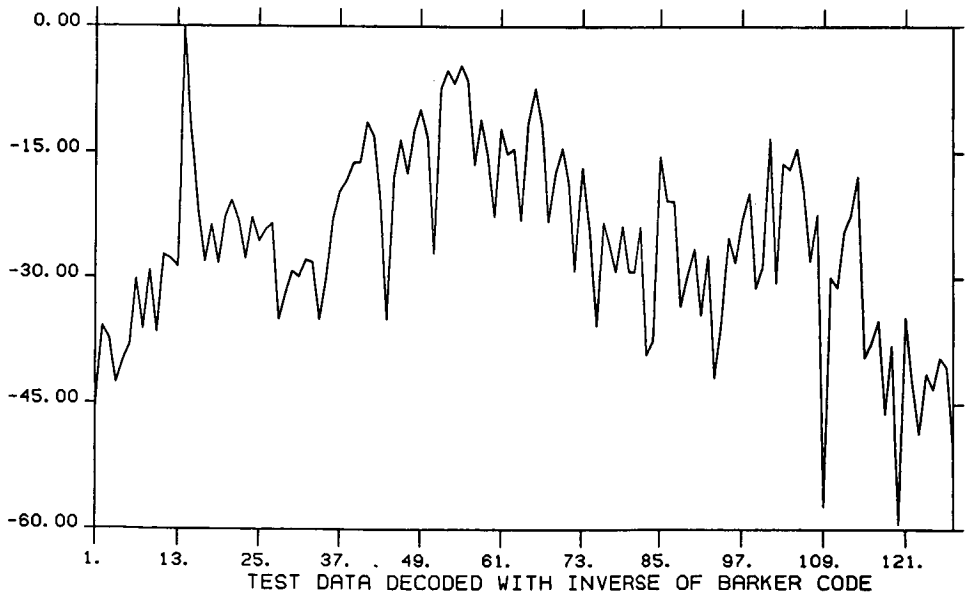


Figure 2.

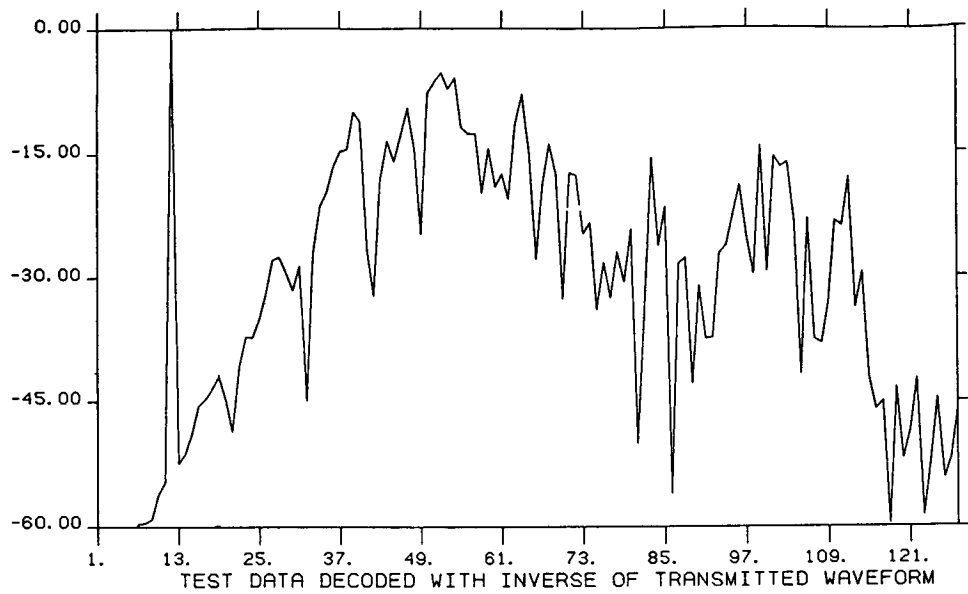


Figure 3.

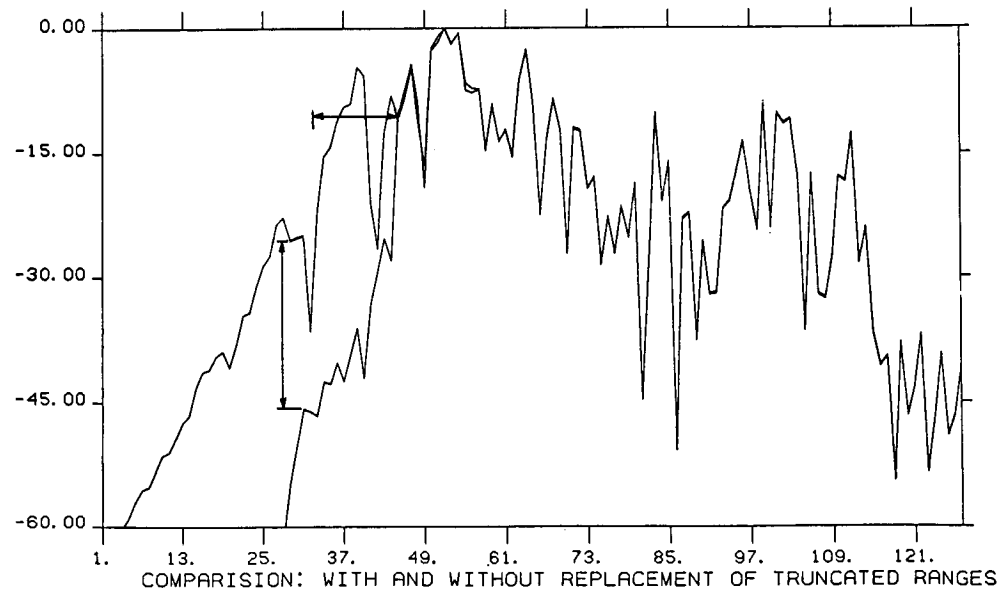


Figure 4.

left are from imperfections in the transmitted waveform. In the forward direction, the sidelobes are not significantly changed since the dominant effect already was the imperfections. The gap between the transmitted and received signals has been partly filled with signal. This has happened because the first range gates of the received signal contain signal from ranges which are truncated by the receiver cutoff. Complete decoding is impossible with either the Barker code or its inverse, but the inverse of the code gives worse response in this respect because the convolving waveform is longer than the code. A comparison of the first two figures reveals some reduction in sidelobe level near the end of the sampled period.

Figure 3 shows the same data decoded with the inverse of the transmitted waveform. Sidelobes are of course completely removed from the transmitted signal. The remaining signal at these delays is due to leakage of the truncated signal to lower altitudes. The range of the leakage has increased due to the increased length of the inverse code. The effect of the Gaussian filter has also been removed, and some signal-to-noise ratio has been lost in doing this. Square pulse matched filters should be used, and then the data will have the ideal triangular shape which is achieved when using a Gaussian filter only by some increase in noise. Finally, the decrease in sidelobes is evident near the end of the sampled time period.

As mentioned before, the leakage due to the truncation of the lower ranges can be reduced. The technique for doing this is explained in detail in Appendix A; briefly, the normal decoding method is used to find the signals at the lower ranges, with sidelobes, of course. The signals from the truncated ranges can then be subtracted away to an accuracy determined by the sidelobe level. The nonmatched decoding technique is used with the result that the range nearest the truncated ranges has leakage about equal to the sidelobe level of the normal decoding method, but the sidelobe levels decrease quickly with increasing range. Figure 4 shows a comparison of the decoding with and without removal of the truncated ranges. The differences in the leakage levels in the direction of decreasing range are similar to the differences expected in the other direction. Using the subtraction technique reduces the leakage by about 20 dB in the lower ranges and it becomes completely insignificant within one pulse width.

PRACTICAL USES OF THE TECHNIQUE

The complementary code pair provides very low sidelobes in many practical circumstances. Two cases where it does not are:

- 1) When the coherence time is short compared to twice the interpulse period. This is usually the case with incoherent scatter.
- 2) When transmitter modulation errors are significant.

As long as the coherence time is longer than the pulse length, we can take any good code and gain a substantial reduction in sidelobes with this technique.

If transmitter modulation errors are a problem but the coherence time is long, then we can use a modification of the technique. The complementary codes are used to modulate the transmitter in the normal way. What is transmitted is somewhat in error, and when we decode, we choose $h_{dc}(t)$ for each complementary code such that the sidelobes of the complementary code are achieved. In other words, we do not try to get rid of the sidelobes, but merely make them what they would have been in the absence of the transmitter errors. As long as transmitter errors are small, this involves a very small correction to the spectral amplitude function, and hence causes no noticeable loss in signal-

to-noise ratio. When the returns from the complementary pair are added, there will be no sidelobes. Since the correction is very small, truncation errors will also be very small.

REFERENCES

- Farley, D. T., On-line data processing techniques for MST radars, Radio Sci., 20, 1177-1184.
 Sulzer, M. P., and R. F. Woodman (1984), Quasi-complementary codes: a new technique for MST radar sounding, Radio Sci., 19, 337-344.

Appendix A

Reducing the Effect of Truncated Ranges

The problem with the truncated ranges that affects the technique described in this paper occurs because given an infinite decoding waveform, the extent of leakage of a range for which the signal is incomplete is infinite. The solution to the problem depends upon the fact that with normal decoding, only those heights that are truncated are affected. The explanation of these two statements requires a detailed examination of the coded waveform.

Figure A1 shows how this waveform can be broken into its component parts. Figure A1a shows the radar signal before decoding. No signal is received to the left of the vertical line because of the receiver cutoff. Figures A1b and A1c show returns from two ranges. When the returns from these two ranges and also from all other ranges are added the signal of Figure A1a is obtained. The signal of Figure A1c is completely to the right of the heavy vertical line, and this means that we have all the information from that range. On the other hand, the signal of Figure 2b is partly to the left of the vertical line, and thus we have only a part of the information from this range. This range is referred to as a truncated range.

Figure A1d shows the waveform used in normal decoding in a position for decoding the lowest nontruncated range. Sidelobes from the truncated ranges are decoded normally, and the lack of information to the left of the vertical line does not affect the decoding of nontruncated ranges. Figure A1e shows the inverse of the Barker code placed in position to decode the same lowest nontruncated range. This waveform extends to the left of the vertical line and thus requires all the information from the truncated ranges in order to reject them completely.

In order to reduce this effect, we decode the first $n-1$ untruncated ranges using normal decoding (n is the length of the code). This waveform contains normal sidelobes from the truncated ranges below and the untruncated ranges above. We recode this signal, which means convolving with the code. Both the wanted signal and the unwanted sidelobes are convolved with the code and thus look like coded signals. Next, we replace the first $n-1$ numbers in the original coded signal with zeros; this is the first $n-1$ samples to the right of the vertical line in Figure A1a. Then, we add to this the recoded waveform from the last step. Finally, inverse decoding is performed on the composite waveform. In the lower ranges, we get sidelobe levels about the same as with normal decoding, but the sidelobes go to zero very quickly as the range increases.

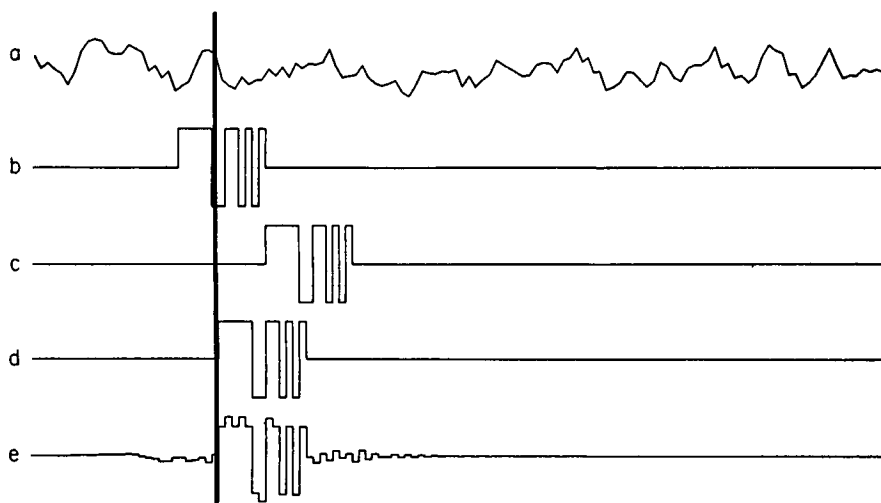


Figure A1.